SEQUENCE LISTING

<110> ROULEAU, Guy A
JOOBER, Richa
BENKELFAT, Chawki

<120> POLYMORPHIC CAG REPEAT-CONTAINING GENE AND USES THEREOF

<130> 2055GG/48747TR

<140> US 09/508,821
<141> 2000-05-26

<150> PCT/CA98/00884
<151> 1998-09-18

<150> CA 2,216,057
<151> 1997-09-19

<160> 11

Sub E7

<170> PatentIn version 3.1

<210> 1
<211> 328
<212> DNA
<213> Mus musculus

<400> 1
agggcagcca ctttccccag cattcccagt ccttccctac ctcctccact tatgccccaa 60

cagtgcaggg tggtgggcag ggggcccact cctacaagag ctgcacagca ccatctgccc 120

agcctcatga taggccgatg agtgccaatg cgaacctggc tccagggcaa cgggtccaga 180

atcttcacgc ttaccagcct ggccgccttg gctacgagca gcagcagcaa gcacttcaag 240

gccgtcacca cacccaggaa acactccact accagaacct cgccaagtac caacactatg 300

gacagcaagg ccagggctac tgtccacc 328

<210> 2
<211> 330
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (59)..(59)
<223> modified base, N = A or G or C or T/U

<400> 2
tccttcccca cctcctccac ctactcctcc tctgtcccag ggtggtgggc agggggccna 60

ctcctataag agttgcacag caccgactgc ccagccccat gacaggccgc tgactgccag 120

ctccagcctg gccccggggc agcgggtcca gaatcttcat gcctaccagt cgggccgcct 180

cagctatgac cagcagcagc agcagcagca gcagcagcag cagcagcagc aagcccttca 240

gagccggcac catgcccagg aaaccctcca ttaccaaaac ctcgccaagt atcagcacta 300

cgggcagcaa ggccagggct actgccagcc 330

<210> 3
<211> 23
<212> DNA
<213> Artificial

<220>
<221> misc_feature
<223> Description of Artificial Sequence: oligonucleotide

<400> 3
ggggcagcgg gtccagaatc ttc 23

<210> 4
<211> 23
<212> DNA
<213> Artificial

<220>
<221> misc_feature
<223> Description of Artificial Sequence: oligonucleotide

<400> 4
agcactacgg gcagcaaggc cag 23

<210> 5
<211> 6022
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (490)..(6021)
<223>

<400> 5
ggatccagca ggcccaaggg gatgagggag cggaaattgc tctgctaaat gcttttgagc 60

tgtcaggaag ggctgggagt gatgggtggg ggacattggg gaggagctgg caatgggcgg 120

gggggggcg ggtagctccc cagtgacctg gcgctgggca gccggttttg cctcccgcat 180

cagtggccgt ccttggcaag actcagctgc aggcgatgtg ggagcgggaa ttacagagca 240

cacctccctg acacagaagt tgtcaatatg cgcacagctg gtggggaggc tcaggcgaag 300

gggggactat taagagctgc gcgggggagc aggcagggtg gggaggtggg tgggagggtg 360

ctttctgagg caaaaggaag tggcccgtct gaatcgctca tcctctgccc cctccctgcc 420

catcctcccc tccctccttc cctccctccc tcccttcctt tttcttttca cagataacca 480

```
gcccgagtc atg cag tct ttt cga gaa agg tgt ggt ttc cat ggc aaa caa     531
          Met Gln Ser Phe Arg Glu Arg Cys Gly Phe His Gly Lys Gln
          1               5                  10

cag aac tac cag cag acc tcg cag gaa aca tca cgc cta gag aat tac     579
Gln Asn Tyr Gln Gln Thr Ser Gln Glu Thr Ser Arg Leu Glu Asn Tyr
15                  20                  25                  30

agg cag ccg agt cag gcc ggg cta agc tgc gac cgg cag cgg ctg ctc     627
Arg Gln Pro Ser Gln Ala Gly Leu Ser Cys Asp Arg Gln Arg Leu Leu
                35                  40                  45

gcc aag gac tat tat aac ccg cag cct tac ccg agc tat gag ggt ggc     675
Ala Lys Asp Tyr Tyr Asn Pro Gln Pro Tyr Pro Ser Tyr Glu Gly Gly
            50                  55                  60

gct ggc acg ccc tct ggc act gca gcc gcg gtg gcc gcc gac aag tac     723
Ala Gly Thr Pro Ser Gly Thr Ala Ala Ala Val Ala Ala Asp Lys Tyr
        65                  70                  75

cac cga ggc agc aag gcc ctg ccc aca cag caa ggc ctg cag ggg agg     771
His Arg Gly Ser Lys Ala Leu Pro Thr Gln Gln Gly Leu Gln Gly Arg
    80                  85                  90

ccg gct ttc cct ggc tac ggc gtc cag gac agc agc ccc tac cca ggc     819
Pro Ala Phe Pro Gly Tyr Gly Val Gln Asp Ser Ser Pro Tyr Pro Gly
95                  100                 105                 110

cgc tat gct ggt gag gag agc ctt cag gct tgg ggg gcc cca cag cca     867
Arg Tyr Ala Gly Glu Glu Ser Leu Gln Ala Trp Gly Ala Pro Gln Pro
                115                 120                 125

cca ccc cca cag ccg cag cca cta cct gca ggg gtg gcc aag tat gat     915
Pro Pro Pro Gln Pro Gln Pro Leu Pro Ala Gly Val Ala Lys Tyr Asp
            130                 135                 140

gag aac ttg atg aaa aag aca gca gtg ccc ccc agc agg cag tat gca     963
Glu Asn Leu Met Lys Lys Thr Ala Val Pro Pro Ser Arg Gln Tyr Ala
        145                 150                 155

gag cag ggc gcc cag gtg ccc ttt cgg act cac tcc ctg cac gtc cag    1011
Glu Gln Gly Ala Gln Val Pro Phe Arg Thr His Ser Leu His Val Gln
    160                 165                 170

cag cca ccg ccg ccc cag cag ccc ctg gca tac ccc aag ctc caa agg    1059
Gln Pro Pro Pro Pro Gln Gln Pro Leu Ala Tyr Pro Lys Leu Gln Arg
175                 180                 185                 190

cag aag ctg cag aac gac att gcc tcc cct ctg ccc ttc ccc cag ggt    1107
Gln Lys Leu Gln Asn Asp Ile Ala Ser Pro Leu Pro Phe Pro Gln Gly
                195                 200                 205

acc cac ttt cct cag cat tcc cag tcc ttc ccc acc tcc tcc acc tac    1155
Thr His Phe Pro Gln His Ser Gln Ser Phe Pro Thr Ser Ser Thr Tyr
            210                 215                 220

tcc tcc tct gtc cag ggt ggt ggg cag ggg gcc cac tcc tat aag agt    1203
Ser Ser Ser Val Gln Gly Gly Gly Gln Gly Ala His Ser Tyr Lys Ser
        225                 230                 235
```

```
tgc aca gca ccg act gcc cag ccc cat gac agg ccg ctg act gcc agc     1251
Cys Thr Ala Pro Thr Ala Gln Pro His Asp Arg Pro Leu Thr Ala Ser
    240                 245                 250

tcc agc ctg gcc ccg ggg cag cgg gtc cag aat ctt cat gcc tac cag     1299
Ser Ser Leu Ala Pro Gly Gln Arg Val Gln Asn Leu His Ala Tyr Gln
255                 260                 265                 270

tcg ggc cgc ctc agc tat gac cag cag cag cag cag cag cag cag cag     1347
Ser Gly Arg Leu Ser Tyr Asp Gln Gln Gln Gln Gln Gln Gln Gln Gln
                275                 280                 285

cag cag cag cag caa gcc ctt cag agc cgg cac cat gcc cag gaa acc     1395
Gln Gln Gln Gln Gln Ala Leu Gln Ser Arg His His Ala Gln Glu Thr
            290                 295                 300

ctc cat tac caa aac ctc gcc aag tat cag cac tac ggg cag caa ggc     1443
Leu His Tyr Gln Asn Leu Ala Lys Tyr Gln His Tyr Gly Gln Gln Gly
        305                 310                 315

cag ggc tac tgc cag ccg gac gca gcc gtc cgg acc cca gag cag tac     1491
Gln Gly Tyr Cys Gln Pro Asp Ala Ala Val Arg Thr Pro Glu Gln Tyr
    320                 325                 330

tac cag acc ttc agc ccc agc tcc agc cac tca ccc gcc cgc tcc gtg     1539
Tyr Gln Thr Phe Ser Pro Ser Ser Ser His Ser Pro Ala Arg Ser Val
335                 340                 345                 350

ggc cgc tca cct tcc tac agt tcc aca ccg tcg ccg ctg atg cca aac     1587
Gly Arg Ser Pro Ser Tyr Ser Ser Thr Pro Ser Pro Leu Met Pro Asn
                355                 360                 365

ctg gag aac ttt ccc tac agc cag cag ccg ctc agc acc ggg gcc ttc     1635
Leu Glu Asn Phe Pro Tyr Ser Gln Gln Pro Leu Ser Thr Gly Ala Phe
            370                 375                 380

ccc gca ggg atc act gac cac agc cac ttc atg ccc ctg ctc aat ccc     1683
Pro Ala Gly Ile Thr Asp His Ser His Phe Met Pro Leu Leu Asn Pro
        385                 390                 395

tcc cca acg gat gcc acc agc tct gtg gac acc cag gct ggc aac tgc     1731
Ser Pro Thr Asp Ala Thr Ser Ser Val Asp Thr Gln Ala Gly Asn Cys
    400                 405                 410

aag ccc ctt cag aag gac aag ctc cct gag aac ctg ctg tcg gat ctc     1779
Lys Pro Leu Gln Lys Asp Lys Leu Pro Glu Asn Leu Leu Ser Asp Leu
415                 420                 425                 430

agc ctg cag agc ctc acg gcg ctg acc tta cag gtg gag aac atc tcc     1827
Ser Leu Gln Ser Leu Thr Ala Leu Thr Leu Gln Val Glu Asn Ile Ser
                435                 440                 445

aac acc gtc cag cag ctg ctg ctc tcc aag gct gct gtg ccg cag aag     1875
Asn Thr Val Gln Gln Leu Leu Leu Ser Lys Ala Ala Val Pro Gln Lys
            450                 455                 460

aaa ggt gtc aag aac ctc gtg tcc agg acc cca gag cag cat aaa agc     1923
Lys Gly Val Lys Asn Leu Val Ser Arg Thr Pro Glu Gln His Lys Ser
        465                 470                 475
```

```
cag cac tgc agc ccc gaa ggg agc ggc tac tca gcc gag ccc gca ggc      1971
Gln His Cys Ser Pro Glu Gly Ser Gly Tyr Ser Ala Glu Pro Ala Gly
    480                 485                 490

aca ccg ctg tca gag ccg ccg agc agc acg cca cag tcc acg cat gcg      2019
Thr Pro Leu Ser Glu Pro Pro Ser Ser Thr Pro Gln Ser Thr His Ala
495                 500                 505                 510

gag ccg cag gag gcc gac tac ctg agc ggc tcc gag gac cca ctg gag      2067
Glu Pro Gln Glu Ala Asp Tyr Leu Ser Gly Ser Glu Asp Pro Leu Glu
                515                 520                 525

cgc agc ttc ctc tac tgc aac cag gcc cgt ggc agc cct gcc agg gtc      2115
Arg Ser Phe Leu Tyr Cys Asn Gln Ala Arg Gly Ser Pro Ala Arg Val
            530                 535                 540

aac agc aac tcg aag gcc aag ccc gag tcc gtg tcc acc tgt tct gtg      2163
Asn Ser Asn Ser Lys Ala Lys Pro Glu Ser Val Ser Thr Cys Ser Val
        545                 550                 555

acc tct cct gac gac atg tcc acc aaa tct gac gac tcc ttc cag agc      2211
Thr Ser Pro Asp Asp Met Ser Thr Lys Ser Asp Asp Ser Phe Gln Ser
    560                 565                 570

cta cac ggc agt ctg ccg ctc gac agc ttc tcc aag ttc gtg gcg ggt      2259
Leu His Gly Ser Leu Pro Leu Asp Ser Phe Ser Lys Phe Val Ala Gly
575                 580                 585                 590

gag cgg gac tgt ccg cgg ctg ctg ctc agc gcc ctg gca cag gag gac      2307
Glu Arg Asp Cys Pro Arg Leu Leu Leu Ser Ala Leu Ala Gln Glu Asp
                595                 600                 605

ctg gcc tcc gag atc ctg ggg ctg cag gaa gcc atc ggt gag aag gcc      2355
Leu Ala Ser Glu Ile Leu Gly Leu Gln Glu Ala Ile Gly Glu Lys Ala
            610                 615                 620

gac aaa gct tgg gct gaa gca ccc agc ctg gtc aag gac agc agc aag      2403
Asp Lys Ala Trp Ala Glu Ala Pro Ser Leu Val Lys Asp Ser Ser Lys
        625                 630                 635

cca ccc ttc tcg ctg gag aac cac agc gcc tgc ctg gac tct gtg gcc      2451
Pro Pro Phe Ser Leu Glu Asn His Ser Ala Cys Leu Asp Ser Val Ala
    640                 645                 650

aag agt gcg tgg ccc cgg cct ggg gag ccg gag gcc ctg ccc gac tcc      2499
Lys Ser Ala Trp Pro Arg Pro Gly Glu Pro Glu Ala Leu Pro Asp Ser
655                 660                 665                 670

ttg cag ctg gac aag ggc ggc aat gcc aag gac ttc agc cca ggg ctg      2547
Leu Gln Leu Asp Lys Gly Gly Asn Ala Lys Asp Phe Ser Pro Gly Leu
                675                 680                 685

ttt gaa gac cct tcc gtg gcc ttc gct acg cct gac ccc aaa aag aca      2595
Phe Glu Asp Pro Ser Val Ala Phe Ala Thr Pro Asp Pro Lys Lys Thr
            690                 695                 700

act ggt cct ctc tcc ttt ggt acc aag ccc acc ctt ggg gtt cct gct      2643
Thr Gly Pro Leu Ser Phe Gly Thr Lys Pro Thr Leu Gly Val Pro Ala
        705                 710                 715
```

```
cca gac ccc act aca gca gct ttt gac tgt ttc ccg gac aca acc gct      2691
Pro Asp Pro Thr Thr Ala Ala Phe Asp Cys Phe Pro Asp Thr Thr Ala
    720                 725                 730

gcc agc tca gcg gac agc gcc aac ccc ttt gcc tgg cca gag gaa aac      2739
Ala Ser Ser Ala Asp Ser Ala Asn Pro Phe Ala Trp Pro Glu Glu Asn
735                 740                 745                 750

ctg ggg gat gct tgt ccc agg tgg gga ttg cac cct ggc gag ctt acc      2787
Leu Gly Asp Ala Cys Pro Arg Trp Gly Leu His Pro Gly Glu Leu Thr
                755                 760                 765

aag ggc ctg gag cag ggt ggg aag gcc tca gat ggc atc agc aaa ggg      2835
Lys Gly Leu Glu Gln Gly Gly Lys Ala Ser Asp Gly Ile Ser Lys Gly
            770                 775                 780

gac acc cat gag gct tcg gcc tgc ctg ggc ttc cag gag gag gac ccc      2883
Asp Thr His Glu Ala Ser Ala Cys Leu Gly Phe Gln Glu Glu Asp Pro
        785                 790                 795

cct ggg gag aag gtg gcc tcg ttg ccc ggg gac ttc aag cag gag gag      2931
Pro Gly Glu Lys Val Ala Ser Leu Pro Gly Asp Phe Lys Gln Glu Glu
    800                 805                 810

gtg ggt ggg gtg aag gag gag gca ggt ggg ctg ctg cag tgc ccc gag      2979
Val Gly Gly Val Lys Glu Glu Ala Gly Gly Leu Leu Gln Cys Pro Glu
815                 820                 825                 830

gtg gcc aag gct gac cgg tgg ctg gag gac agc cgg cac tgc tgt tcc      3027
Val Ala Lys Ala Asp Arg Trp Leu Glu Asp Ser Arg His Cys Cys Ser
                835                 840                 845

acc gcc gac ttc ggg gac ctc cca ctg ctg cca ccc acc agc agg aag      3075
Thr Ala Asp Phe Gly Asp Leu Pro Leu Leu Pro Pro Thr Ser Arg Lys
            850                 855                 860

gag gac ctg gaa gct gag gag gag tac tcc tcc cta tgt gag ctc ctg      3123
Glu Asp Leu Glu Ala Glu Glu Glu Tyr Ser Ser Leu Cys Glu Leu Leu
        865                 870                 875

ggc agc ccc gag cag agg cct ggc atg cag gac ccg ctg tca ccc aag      3171
Gly Ser Pro Glu Gln Arg Pro Gly Met Gln Asp Pro Leu Ser Pro Lys
    880                 885                 890

gcc cca ctc atc tgc acc aag gag gag gtg gag gag gtg ctg gac tcc      3219
Ala Pro Leu Ile Cys Thr Lys Glu Glu Val Glu Glu Val Leu Asp Ser
895                 900                 905                 910

aag gcc ggc tgg ggc tct ccg tgc cac ctc tca ggg gag tcc gtc atc      3267
Lys Ala Gly Trp Gly Ser Pro Cys His Leu Ser Gly Glu Ser Val Ile
                915                 920                 925

ctg ctg ggc cct aca gtg ggc acc gag tca aag gtc cag agc tgg ttt      3315
Leu Leu Gly Pro Thr Val Gly Thr Glu Ser Lys Val Gln Ser Trp Phe
            930                 935                 940

gag tcc tct ctg tca cac atg aag cca ggt gaa gag ggg cct gat ggg      3363
Glu Ser Ser Leu Ser His Met Lys Pro Gly Glu Glu Gly Pro Asp Gly
        945                 950                 955
```

```
gag cga gct cca ggg gat tcc acc acc tcg gac gcc tct ctg gcc cag      3411
Glu Arg Ala Pro Gly Asp Ser Thr Thr Ser Asp Ala Ser Leu Ala Gln
    960                 965                 970

aag ccc aac aag cct gct gtg ccc gag gcg ccc atc gca aag aaa gag      3459
Lys Pro Asn Lys Pro Ala Val Pro Glu Ala Pro Ile Ala Lys Lys Glu
975                 980                 985                 990

cct gtg cca cgg ggc aaa agc tta cgg agc  cgt cgg gtg cac cgg  ggg    3507
Pro Val Pro Arg Gly Lys Ser Leu Arg Ser  Arg Arg Val His Arg  Gly
                995                1000                1005

ctg ccc gag gcc  gag gac tcc cca tgc  agg gca cca gtg ctg  ccc       3552
Leu Pro Glu Ala  Glu Asp Ser Pro Cys  Arg Ala Pro Val Leu  Pro
            1010                1015                1020

aaa gac ctc ttg  ctc cct gaa tcc tgc  aca ggg ccc ccc cag  gga       3597
Lys Asp Leu Leu  Leu Pro Glu Ser Cys  Thr Gly Pro Pro Gln  Gly
            1025                1030                1035

cag atg gaa ggg  gct gga gcc cca ggc  cgg ggg gcc tcg gaa  ggg       3642
Gln Met Glu Gly  Ala Gly Ala Pro Gly  Arg Gly Ala Ser Glu  Gly
            1040                1045                1050

ctc ccc agg atg  tgt act cgt tct ctc  acg gcc ctg agt gag  ccc       3687
Leu Pro Arg Met  Cys Thr Arg Ser Leu  Thr Ala Leu Ser Glu  Pro
            1055                1060                1065

cgc acg ccc gga  ccc cca ggc ctg acc  acc acc cct gca ccc  cca       3732
Arg Thr Pro Gly  Pro Pro Gly Leu Thr  Thr Thr Pro Ala Pro  Pro
            1070                1075                1080

gac aaa ctg ggg  ggc aag cag cga gcc  gcc ttc aag tcg ggc  aag       3777
Asp Lys Leu Gly  Gly Lys Gln Arg Ala  Ala Phe Lys Ser Gly  Lys
            1085                1090                1095

cgg gtg ggg aag  ccc tca ccc aag gct  gcc tcc agc ccc agc  aac       3822
Arg Val Gly Lys  Pro Ser Pro Lys Ala  Ala Ser Ser Pro Ser  Asn
            1100                1105                1110

ccg gcc gcc ctg  cct gtg gcc tcc gac  agc agc ccg atg ggc  tcc       3867
Pro Ala Ala Leu  Pro Val Ala Ser Asp  Ser Ser Pro Met Gly  Ser
            1115                1120                1125

aag acc aag gag  aca gac tca ccc agc  acg cct ggc aag gac  cag       3912
Lys Thr Lys Glu  Thr Asp Ser Pro Ser  Thr Pro Gly Lys Asp  Gln
            1130                1135                1140

cgc tcc atg atc  ctt cgg tca cgc acc  aaa acc cag gag atc  ttc       3957
Arg Ser Met Ile  Leu Arg Ser Arg Thr  Lys Thr Gln Glu Ile  Phe
            1145                1150                1155

cac tcc aag cgg  cgg agg ccc tct gag  ggc cgg ctc ccc aac  tgc       4002
His Ser Lys Arg  Arg Arg Pro Ser Glu  Gly Arg Leu Pro Asn  Cys
            1160                1165                1170

cgt gcc acc aag  aag ctc ctc gac aac  agc cac ttg ccc gcc  aca       4047
Arg Ala Thr Lys  Lys Leu Leu Asp Asn  Ser His Leu Pro Ala  Thr
            1175                1180                1185
```

```
ttc aag gtc tcc  agc agc ccc cag aag  gag ggc agg gtg agc  cag      4092
Phe Lys Val Ser  Ser Ser Pro Gln Lys  Glu Gly Arg Val Ser  Gln
            1190                 1195                 1200

cgg gca agg gtc  ccc aaa cct ggt gca  ggc agc aag ctc tct  gac      4137
Arg Ala Arg Val  Pro Lys Pro Gly Ala  Gly Ser Lys Leu Ser  Asp
            1205                 1210                 1215

cgg ccc ctc cat  gcg ctc aaa agg aag  tcg gcc ttc atg gcg  ccg      4182
Arg Pro Leu His  Ala Leu Lys Arg Lys  Ser Ala Phe Met Ala  Pro
            1220                 1225                 1230

gtc ccc acc aag  aag cgg aac ctg gtc  ttg cgg cac ggc agc  agc      4227
Val Pro Thr Lys  Lys Arg Asn Leu Val  Leu Arg His Gly Ser  Ser
            1235                 1240                 1245

agc agc agc aac  gcc agt gca atg ggg  gag atg gga agg agg  aga      4272
Ser Ser Ser Asn  Ala Ser Ala Met Gly  Glu Met Gly Arg Arg  Arg
            1250                 1255                 1260

ggc ctg agg gtt  ccc cca ccc tct tca  aga gga tgt ctt ctc  cca      4317
Gly Leu Arg Val  Pro Pro Pro Ser Ser  Arg Gly Cys Leu Leu  Pro
            1265                 1270                 1275

aga aag cca agc  cca cca agg gca atg  gcg agc ctg cca caa  agc      4362
Arg Lys Pro Ser  Pro Pro Arg Ala Met  Ala Ser Leu Pro Gln  Ser
            1280                 1285                 1290

tcc cac ccc cgg  aga ccc cca ttc ctg  cct caa gct cgc ctc  tcg      4407
Ser His Pro Arg  Arg Pro Pro Phe Leu  Pro Gln Ala Arg Leu  Ser
            1295                 1300                 1305

gca gcc ttc cag  ggg gcc atg aag acc  aag gtg ctg cca ccc  cgg      4452
Ala Ala Phe Gln  Gly Ala Met Lys Thr  Lys Val Leu Pro Pro  Arg
            1310                 1315                 1320

aag ggc cgg ggc  ctg aag ctg gaa gcc  atc gtg cag aag atc  acc      4497
Lys Gly Arg Gly  Leu Lys Leu Glu Ala  Ile Val Gln Lys Ile  Thr
            1325                 1330                 1335

tcg ccc agc ctc  aag aag ttc gca tgt  aaa gcg cca ggg gcc  tct      4542
Ser Pro Ser Leu  Lys Lys Phe Ala Cys  Lys Ala Pro Gly Ala  Ser
            1340                 1345                 1350

cct ggt aat cct  ctg agc cca tcc ctt  tcc gac aaa gac cgt  ggg      4587
Pro Gly Asn Pro  Leu Ser Pro Ser Leu  Ser Asp Lys Asp Arg  Gly
            1355                 1360                 1365

ctc aag ggt gct  ggg ggc agc cca gtg  ggg gtg gaa gaa ggc  ctg      4632
Leu Lys Gly Ala  Gly Gly Ser Pro Val  Gly Val Glu Glu Gly  Leu
            1370                 1375                 1380

gta aat gtg ggc  acc ggg cag aag ctc  cca act tct ggg gct  gat      4677
Val Asn Val Gly  Thr Gly Gln Lys Leu  Pro Thr Ser Gly Ala  Asp
            1385                 1390                 1395

ccg tta tgc aga  aat cca acc aac aga  tcc tta aaa ggc aaa  ctc      4722
Pro Leu Cys Arg  Asn Pro Thr Asn Arg  Ser Leu Lys Gly Lys  Leu
            1400                 1405                 1410
```

```
atg aac agt aag  aaa ctg tct tct act  gac tgt ttc aaa acc  gag      4767
Met Asn Ser Lys  Lys Leu Ser Ser Thr  Asp Cys Phe Lys Thr  Glu
            1415                 1420                 1425

gcc ttc aca tcc  ccg gag gcc ctg cag  cct ggg ggg act gcc  ctg      4812
Ala Phe Thr Ser  Pro Glu Ala Leu Gln  Pro Gly Gly Thr Ala  Leu
            1430                 1435                 1440

gcg cct aag aag  agg agc cgg aaa ggc  cgg gca ggg gcc cat  gga      4857
Ala Pro Lys Lys  Arg Ser Arg Lys Gly  Arg Ala Gly Ala His  Gly
            1445                 1450                 1455

ctc tcc aaa ggc  ccg ctg gag aag cgg  ccc tat ctt ggc ccg  gct      4902
Leu Ser Lys Gly  Pro Leu Glu Lys Arg  Pro Tyr Leu Gly Pro  Ala
            1460                 1465                 1470

ctg ctc ctg act  ccc cga gac agg gcc  agt ggc aca caa ggg  gcc      4947
Leu Leu Leu Thr  Pro Arg Asp Arg Ala  Ser Gly Thr Gln Gly  Ala
            1475                 1480                 1485

agt gag gac aac  tct ggt gga gga ggc  aag aag cca aag atg  gag      4992
Ser Glu Asp Asn  Ser Gly Gly Gly Gly  Lys Lys Pro Lys Met  Glu
            1490                 1495                 1500

gag ctg ggc cct  gcc tcc cag ccc ccg  gag ggc agg ccc tgc  cag      5037
Glu Leu Gly Pro  Ala Ser Gln Pro Pro  Glu Gly Arg Pro Cys  Gln
            1505                 1510                 1515

ccc cag aca agg  gca cag aaa cag cca  ggc cac acc aac tac  agc      5082
Pro Gln Thr Arg  Ala Gln Lys Gln Pro  Gly His Thr Asn Tyr  Ser
            1520                 1525                 1530

agc tat tcc aag  cgg aag cgc ctc act  cgg ggc cgg gcc aag  aac      5127
Ser Tyr Ser Lys  Arg Lys Arg Leu Thr  Arg Gly Arg Ala Lys  Asn
            1535                 1540                 1545

acc acc tct tca  ccc tgt aag ggg cgt  gcc aag cga cga cga  caa      5172
Thr Thr Ser Ser  Pro Cys Lys Gly Arg  Ala Lys Arg Arg Arg  Gln
            1550                 1555                 1560

cag cag gtg ctg  ccc ctg gat ccc gca  gag cct gaa atc cgc  ctc      5217
Gln Gln Val Leu  Pro Leu Asp Pro Ala  Glu Pro Glu Ile Arg  Leu
            1565                 1570                 1575

aag tac att tcc  tct tgc aag cgg ctg  agg tca gac agc cgg  acc      5262
Lys Tyr Ile Ser  Ser Cys Lys Arg Leu  Arg Ser Asp Ser Arg  Thr
            1580                 1585                 1590

ccc gcc ttc tca  ccc ttc gtg cgg gtg  gag aag cga gac gcg  ttc      5307
Pro Ala Phe Ser  Pro Phe Val Arg Val  Glu Lys Arg Asp Ala  Phe
            1595                 1600                 1605

acc acc ata tgc  act gtt gtc aac tcc  cct gga gat gcg ccc  aag      5352
Thr Thr Ile Cys  Thr Val Val Asn Ser  Pro Gly Asp Ala Pro  Lys
            1610                 1615                 1620

ccc cac agg aag  cct tcc tcc tct gcc  tcc tct tcc tca tcc  tcg      5397
Pro His Arg Lys  Pro Ser Ser Ser Ala  Ser Ser Ser Ser Ser  Ser
            1625                 1630                 1635
```

```
tcc tcg ttc tcc  ttg gat gca gcc ggg  gcc tcc ctg gcc aca  ctc      5442
Ser Ser Phe Ser  Leu Asp Ala Ala Gly  Ala Ser Leu Ala Thr  Leu
            1640                 1645                 1650

cct gga ggc tcc  atc ctg cag ccg cgg  ccc tcc ttg ccc ctc  tcc      5487
Pro Gly Gly Ser  Ile Leu Gln Pro Arg  Pro Ser Leu Pro Leu  Ser
            1655                 1660                 1665

tcc acg atg cac  ttg ggg cct gtg gtt  tcc aag gcc ctg agt  acc      5532
Ser Thr Met His  Leu Gly Pro Val Val  Ser Lys Ala Leu Ser  Thr
            1670                 1675                 1680

tct tgc ctt gtt  tgc tgc ctc tgc caa  aac ccg gcc aac ttc  aag      5577
Ser Cys Leu Val  Cys Cys Leu Cys Gln  Asn Pro Ala Asn Phe  Lys
            1685                 1690                 1695

gac ctt ggg gac  ctc tgt ggg ccc tac  tac cct gaa cac tgc  ctc      5622
Asp Leu Gly Asp  Leu Cys Gly Pro Tyr  Tyr Pro Glu His Cys  Leu
            1700                 1705                 1710

ccc aaa aag aag  cca aaa ctc aag gag  aag gtg cgg cca gaa  ggc      5667
Pro Lys Lys Lys  Pro Lys Leu Lys Glu  Lys Val Arg Pro Glu  Gly
            1715                 1720                 1725

acc tgt gag gag  gcc tcg ctg ccg ctt  gag aga aca ctc aaa  ggt      5712
Thr Cys Glu Glu  Ala Ser Leu Pro Leu  Glu Arg Thr Leu Lys  Gly
            1730                 1735                 1740

ccc gag tgt gca  gct gcc gcc act gcc  ggg aag ccc ccc agg  tga      5757
Pro Glu Cys Ala  Ala Ala Ala Thr Ala  Gly Lys Pro Pro Arg
            1745                 1750                 1755

cgg ccc agc tga ccc ggc  caa gca ggg ccc act  gcg cac cag tgc      5802
Arg Pro Ser     Pro Gly  Gln Ala Gly Pro Thr  Ala His Gln Cys
                    1760                 1765

ccg  ggg cct gtc ccg gag  gct gca gag ctg cta  ctg ctg tga tgg      5847
Pro  Gly Pro Val Pro Glu  Ala Ala Glu Leu Leu  Leu Leu     Trp
1770                 1775                 1780

ccg gga  gga tgg ggg cga gga  ggc agc ccc agc cga  caa ggg tcg      5892
Pro Gly  Gly Trp Gly Arg Gly  Gly Ser Pro Ser Arg  Gln Gly Ser
    1785                 1790                 1795

caa aca  tga gtg cag caa gga ggc  tcc ggc aga gcc cgg  cgg gga      5937
Gln Thr      Val Gln Gln Gly Gly  Ser Gly Arg Ala Arg  Arg Gly
    1800                 1805                 1810

ggc cca gga  gca ctg ggt gca tga ggc  ctg tgc cgt gtg gac  cgg      5982
Gly Pro Gly  Ala Leu Gly Ala     Gly  Leu Cys Arg Val Asp  Arg
        1815                 1820                 1825

cgg cgt cta cct  ggt ggc cgg gaa gct  ctt tgg gct gca g           6022
Arg Arg Leu Pro  Gly Gly Arg Glu Ala  Leu Trp Ala Ala
            1830                 1835
```

<210> 6
<211> 1755
<212> PRT
<213> Homo sapiens

<400> 6

Met Gln Ser Phe Arg Glu Arg Cys Gly Phe His Gly Lys Gln Gln Asn
1 5 10 15

Tyr Gln Gln Thr Ser Gln Glu Thr Ser Arg Leu Glu Asn Tyr Arg Gln
20 25 30

Pro Ser Gln Ala Gly Leu Ser Cys Asp Arg Gln Arg Leu Leu Ala Lys
35 40 45

Asp Tyr Tyr Asn Pro Gln Pro Tyr Pro Ser Tyr Glu Gly Gly Ala Gly
50 55 60

Thr Pro Ser Gly Thr Ala Ala Ala Val Ala Ala Asp Lys Tyr His Arg
65 70 75 80

Gly Ser Lys Ala Leu Pro Thr Gln Gln Gly Leu Gln Gly Arg Pro Ala
85 90 95

Phe Pro Gly Tyr Gly Val Gln Asp Ser Ser Pro Tyr Pro Gly Arg Tyr
100 105 110

Ala Gly Glu Glu Ser Leu Gln Ala Trp Gly Ala Pro Gln Pro Pro Pro
115 120 125

Pro Gln Pro Gln Pro Leu Pro Ala Gly Val Ala Lys Tyr Asp Glu Asn
130 135 140

Leu Met Lys Lys Thr Ala Val Pro Pro Ser Arg Gln Tyr Ala Glu Gln
145 150 155 160

Gly Ala Gln Val Pro Phe Arg Thr His Ser Leu His Val Gln Gln Pro
165 170 175

Pro Pro Pro Gln Gln Pro Leu Ala Tyr Pro Lys Leu Gln Arg Gln Lys
180 185 190

Leu Gln Asn Asp Ile Ala Ser Pro Leu Pro Phe Pro Gln Gly Thr His
195 200 205

Phe Pro Gln His Ser Gln Ser Phe Pro Thr Ser Ser Thr Tyr Ser Ser
210 215 220

Ser Val Gln Gly Gly Gly Gln Gly Ala His Ser Tyr Lys Ser Cys Thr
225 230 235 240

Ala Pro Thr Ala Gln Pro His Asp Arg Pro Leu Thr Ala Ser Ser Ser
245 250 255

Leu Ala Pro Gly Gln Arg Val Gln Asn Leu His Ala Tyr Gln Ser Gly
260 265 270

Arg Leu Ser Tyr Asp Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
275 280 285

Gln Gln Gln Ala Leu Gln Ser Arg His His Ala Gln Glu Thr Leu His
290 295 300

Tyr Gln Asn Leu Ala Lys Tyr Gln His Tyr Gly Gln Gln Gly Gln Gly
305 310 315 320

Tyr Cys Gln Pro Asp Ala Ala Val Arg Thr Pro Glu Gln Tyr Tyr Gln
325 330 335

Thr Phe Ser Pro Ser Ser Ser His Ser Pro Ala Arg Ser Val Gly Arg
340 345 350

Ser Pro Ser Tyr Ser Ser Thr Pro Ser Pro Leu Met Pro Asn Leu Glu
355 360 365

Asn Phe Pro Tyr Ser Gln Gln Pro Leu Ser Thr Gly Ala Phe Pro Ala
370 375 380

Gly Ile Thr Asp His Ser His Phe Met Pro Leu Leu Asn Pro Ser Pro
385 390 395 400

Thr Asp Ala Thr Ser Ser Val Asp Thr Gln Ala Gly Asn Cys Lys Pro
405 410 415

Leu Gln Lys Asp Lys Leu Pro Glu Asn Leu Leu Ser Asp Leu Ser Leu
420 425 430

Gln Ser Leu Thr Ala Leu Thr Leu Gln Val Glu Asn Ile Ser Asn Thr
435 440 445

Val Gln Gln Leu Leu Leu Ser Lys Ala Ala Val Pro Gln Lys Lys Gly
450 455 460

Val Lys Asn Leu Val Ser Arg Thr Pro Glu Gln His Lys Ser Gln His
465 470 475 480

Cys Ser Pro Glu Gly Ser Gly Tyr Ser Ala Glu Pro Ala Gly Thr Pro
485 490 495

Leu Ser Glu Pro Pro Ser Ser Thr Pro Gln Ser Thr His Ala Glu Pro
500 505 510

Gln Glu Ala Asp Tyr Leu Ser Gly Ser Glu Asp Pro Leu Glu Arg Ser
515 520 525

Phe Leu Tyr Cys Asn Gln Ala Arg Gly Ser Pro Ala Arg Val Asn Ser
530 535 540

Asn Ser Lys Ala Lys Pro Glu Ser Val Ser Thr Cys Ser Val Thr Ser
545 550 555 560

Pro Asp Asp Met Ser Thr Lys Ser Asp Asp Ser Phe Gln Ser Leu His
565 570 575

Gly Ser Leu Pro Leu Asp Ser Phe Ser Lys Phe Val Ala Gly Glu Arg
580 585 590

Asp Cys Pro Arg Leu Leu Leu Ser Ala Leu Ala Gln Glu Asp Leu Ala
595 600 605

Ser Glu Ile Leu Gly Leu Gln Glu Ala Ile Gly Glu Lys Ala Asp Lys
610 615 620

Ala Trp Ala Glu Ala Pro Ser Leu Val Lys Asp Ser Ser Lys Pro Pro
625 630 635 640

Phe Ser Leu Glu Asn His Ser Ala Cys Leu Asp Ser Val Ala Lys Ser
645 650 655

Ala Trp Pro Arg Pro Gly Glu Pro Glu Ala Leu Pro Asp Ser Leu Gln
660 665 670

Leu Asp Lys Gly Gly Asn Ala Lys Asp Phe Ser Pro Gly Leu Phe Glu
675 680 685

Asp Pro Ser Val Ala Phe Ala Thr Pro Asp Pro Lys Lys Thr Thr Gly
690 695 700

Pro Leu Ser Phe Gly Thr Lys Pro Thr Leu Gly Val Pro Ala Pro Asp
705 710 715 720

Pro Thr Thr Ala Ala Phe Asp Cys Phe Pro Asp Thr Thr Ala Ala Ser
725 730 735

Ser Ala Asp Ser Ala Asn Pro Phe Ala Trp Pro Glu Glu Asn Leu Gly
740 745 750

Asp Ala Cys Pro Arg Trp Gly Leu His Pro Gly Glu Leu Thr Lys Gly
755 760 765

Leu Glu Gln Gly Gly Lys Ala Ser Asp Gly Ile Ser Lys Gly Asp Thr
770 775 780

His Glu Ala Ser Ala Cys Leu Gly Phe Gln Glu Glu Asp Pro Pro Gly
785 790 795 800

Glu Lys Val Ala Ser Leu Pro Gly Asp Phe Lys Gln Glu Glu Val Gly
805 810 815

Gly Val Lys Glu Glu Ala Gly Gly Leu Leu Gln Cys Pro Glu Val Ala
820 825 830

Lys Ala Asp Arg Trp Leu Glu Asp Ser Arg His Cys Cys Ser Thr Ala
835 840 845

Asp Phe Gly Asp Leu Pro Leu Leu Pro Pro Thr Ser Arg Lys Glu Asp
850 855 860

Leu Glu Ala Glu Glu Glu Tyr Ser Ser Leu Cys Glu Leu Leu Gly Ser
865 870 875 880

Pro Glu Gln Arg Pro Gly Met Gln Asp Pro Leu Ser Pro Lys Ala Pro
885 890 895

Leu Ile Cys Thr Lys Glu Glu Val Glu Glu Val Leu Asp Ser Lys Ala
900 905 910

Gly Trp Gly Ser Pro Cys His Leu Ser Gly Glu Ser Val Ile Leu Leu
915 920 925

Gly Pro Thr Val Gly Thr Glu Ser Lys Val Gln Ser Trp Phe Glu Ser
930 935 940

```
Ser Leu Ser His Met Lys Pro Gly Glu Glu Gly Pro Asp Gly Glu Arg
945                 950                 955                 960

Ala Pro Gly Asp Ser Thr Thr Ser Asp Ala Ser Leu Ala Gln Lys Pro
                965                 970                 975

Asn Lys Pro Ala Val Pro Glu Ala Pro Ile Ala Lys Lys Glu Pro Val
            980                 985                 990

Pro Arg Gly Lys Ser Leu Arg Ser  Arg Arg Val His Arg  Gly Leu Pro
        995                 1000                1005

Glu Ala  Glu Asp Ser Pro Cys  Arg Ala Pro Val Leu  Pro Lys Asp
    1010                1015                1020

Leu Leu  Leu Pro Glu Ser Cys  Thr Gly Pro Pro Gln  Gly Gln Met
    1025                1030                1035

Glu Gly  Ala Gly Ala Pro Gly  Arg Gly Ala Ser Glu  Gly Leu Pro
    1040                1045                1050

Arg Met  Cys Thr Arg Ser Leu  Thr Ala Leu Ser Glu  Pro Arg Thr
    1055                1060                1065

Pro Gly  Pro Pro Gly Leu Thr  Thr Thr Pro Ala Pro  Pro Asp Lys
    1070                1075                1080

Leu Gly  Gly Lys Gln Arg Ala  Ala Phe Lys Ser Gly  Lys Arg Val
    1085                1090                1095

Gly Lys  Pro Ser Pro Lys Ala  Ala Ser Ser Pro Ser  Asn Pro Ala
    1100                1105                1110

Ala Leu  Pro Val Ala Ser Asp  Ser Ser Pro Met Gly  Ser Lys Thr
    1115                1120                1125

Lys Glu  Thr Asp Ser Pro Ser  Thr Pro Gly Lys Asp  Gln Arg Ser
    1130                1135                1140

Met Ile  Leu Arg Ser Arg Thr  Lys Thr Gln Glu Ile  Phe His Ser
    1145                1150                1155

Lys Arg  Arg Arg Pro Ser Glu  Gly Arg Leu Pro Asn  Cys Arg Ala
    1160                1165                1170

Thr Lys  Lys Leu Leu Asp Asn  Ser His Leu Pro Ala  Thr Phe Lys
    1175                1180                1185

Val Ser  Ser Ser Pro Gln Lys  Glu Gly Arg Val Ser  Gln Arg Ala
    1190                1195                1200

Arg Val  Pro Lys Pro Gly Ala  Gly Ser Lys Leu Ser  Asp Arg Pro
    1205                1210                1215

Leu His  Ala Leu Lys Arg Lys  Ser Ala Phe Met Ala  Pro Val Pro
    1220                1225                1230

Thr Lys  Lys Arg Asn Leu Val  Leu Arg His Gly Ser  Ser Ser Ser
    1235                1240                1245
```

Ser Asn Ala Ser Ala Met Gly Glu Met Gly Arg Arg Arg Gly Leu
1250 1255 1260

Arg Val Pro Pro Pro Ser Ser Arg Gly Cys Leu Leu Pro Arg Lys
1265 1270 1275

Pro Ser Pro Pro Arg Ala Met Ala Ser Leu Pro Gln Ser Ser His
1280 1285 1290

Pro Arg Arg Pro Pro Phe Leu Pro Gln Ala Arg Leu Ser Ala Ala
1295 1300 1305

Phe Gln Gly Ala Met Lys Thr Lys Val Leu Pro Pro Arg Lys Gly
1310 1315 1320

Arg Gly Leu Lys Leu Glu Ala Ile Val Gln Lys Ile Thr Ser Pro
1325 1330 1335

Ser Leu Lys Lys Phe Ala Cys Lys Ala Pro Gly Ala Ser Pro Gly
1340 1345 1350

Asn Pro Leu Ser Pro Ser Leu Ser Asp Lys Asp Arg Gly Leu Lys
1355 1360 1365

Gly Ala Gly Gly Ser Pro Val Gly Val Glu Glu Gly Leu Val Asn
1370 1375 1380

Val Gly Thr Gly Gln Lys Leu Pro Thr Ser Gly Ala Asp Pro Leu
1385 1390 1395

Cys Arg Asn Pro Thr Asn Arg Ser Leu Lys Gly Lys Leu Met Asn
1400 1405 1410

Ser Lys Lys Leu Ser Ser Thr Asp Cys Phe Lys Thr Glu Ala Phe
1415 1420 1425

Thr Ser Pro Glu Ala Leu Gln Pro Gly Gly Thr Ala Leu Ala Pro
1430 1435 1440

Lys Lys Arg Ser Arg Lys Gly Arg Ala Gly Ala His Gly Leu Ser
1445 1450 1455

Lys Gly Pro Leu Glu Lys Arg Pro Tyr Leu Gly Pro Ala Leu Leu
1460 1465 1470

Leu Thr Pro Arg Asp Arg Ala Ser Gly Thr Gln Gly Ala Ser Glu
1475 1480 1485

Asp Asn Ser Gly Gly Gly Gly Lys Lys Pro Lys Met Glu Glu Leu
1490 1495 1500

Gly Pro Ala Ser Gln Pro Pro Glu Gly Arg Pro Cys Gln Pro Gln
1505 1510 1515

Thr Arg Ala Gln Lys Gln Pro Gly His Thr Asn Tyr Ser Ser Tyr
1520 1525 1530

Ser Lys Arg Lys Arg Leu Thr Arg Gly Arg Ala Lys Asn Thr Thr
1535 1540 1545

```
Ser Ser  Pro Cys Lys Gly Arg  Ala Lys Arg Arg Arg  Gln Gln Gln
    1550                 1555                 1560

Val Leu  Pro Leu Asp Pro Ala  Glu Pro Glu Ile Arg  Leu Lys Tyr
    1565                 1570                 1575

Ile Ser  Ser Cys Lys Arg Leu  Arg Ser Asp Ser Arg  Thr Pro Ala
    1580                 1585                 1590

Phe Ser  Pro Phe Val Arg Val  Glu Lys Arg Asp Ala  Phe Thr Thr
    1595                 1600                 1605

Ile Cys  Thr Val Val Asn Ser  Pro Gly Asp Ala Pro  Lys Pro His
    1610                 1615                 1620

Arg Lys  Pro Ser Ser Ser Ala  Ser Ser Ser Ser Ser  Ser Ser Ser
    1625                 1630                 1635

Phe Ser  Leu Asp Ala Ala Gly  Ala Ser Leu Ala Thr  Leu Pro Gly
    1640                 1645                 1650

Gly Ser  Ile Leu Gln Pro Arg  Pro Ser Leu Pro Leu  Ser Ser Thr
    1655                 1660                 1665

Met His  Leu Gly Pro Val Val  Ser Lys Ala Leu Ser  Thr Ser Cys
    1670                 1675                 1680

Leu Val  Cys Cys Leu Cys Gln  Asn Pro Ala Asn Phe  Lys Asp Leu
    1685                 1690                 1695

Gly Asp  Leu Cys Gly Pro Tyr  Tyr Pro Glu His Cys  Leu Pro Lys
    1700                 1705                 1710

Lys Lys  Pro Lys Leu Lys Glu  Lys Val Arg Pro Glu  Gly Thr Cys
    1715                 1720                 1725

Glu Glu  Ala Ser Leu Pro Leu  Glu Arg Thr Leu Lys  Gly Pro Glu
    1730                 1735                 1740

Cys Ala  Ala Ala Ala Thr Ala  Gly Lys Pro Pro Arg
    1745                 1750                 1755


<210>  7
<211>  24
<212>  PRT
<213>  Homo sapiens

<400>  7

Pro Gly Gln Ala Gly Pro Thr Ala His Gln Cys Pro Gly Pro Val Pro
1               5                  10                  15

Glu Ala Ala Glu Leu Leu Leu Leu
            20


<210>  8
<211>  18
<212>  PRT
<213>  Homo sapiens
```

<400> 8

Trp Pro Gly Gly Trp Gly Arg Gly Gly Ser Pro Ser Arg Gln Gly Ser
1 5 10 15

Gln Thr

<210> 9
<211> 19
<212> PRT
<213> Homo sapiens

<400> 9

Val Gln Gln Gly Gly Ser Gly Arg Ala Arg Arg Gly Gly Pro Gly Ala
1 5 10 15

Leu Gly Ala

Sub E7

<210> 10
<211> 20
<212> PRT
<213> Homo sapiens

<400> 10

Gly Leu Cys Arg Val Asp Arg Arg Arg Leu Pro Gly Gly Arg Glu Ala
1 5 10 15

Leu Trp Ala Ala
20

<210> 11
<211> 22
<212> DNA
<213> Artificial

<220>
<221> misc_feature
<223> Description of Artificial Sequence: oligonucleotide

<400> 11
tggccttgct gcccgtagtg ct 22